Review problems

The problems in the final exam could be significantly different from the following ones. In the exam, everyone is allowed to bring a two sided review sheet with size 8×11. We will provide normal and t tables, and the table of distributions with PDF, mean and variance, on page 433 of the textbook.

FINAL EXAM: GR109, ROSENFELD, May 09, 2009 beginning at 02:00pm.

1. Let $X_1, X_2, \ldots, X_n$ be i.i.d. $N(0, \tau)$ with unknown $\tau > 0$.
(a). What's the log-likelihood function?
(b). What's the maximum likelihood estimator of τ?
(c). What is the Fisher information $I(\tau)$?
(d). Let $\widehat{\tau} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$ be an estimator of τ. Is it unbiased? What is its variance?
(e). What is the asymptotic distribution of $\widehat{\tau}$?
(f). Give the form of an approximate 95% confidence interval for τ?
(g). Let $\frac{1}{n+2}\sum_{i=1}^{n} X_i^2$ be another estimator of τ. Is it unbiased? What is its variance?
(h). Let the prior $f(\tau) = \tau$. What's the posterior density and posterior mean? Find an approximate 95% posterior interval for τ.
(i). Compare the mean square error of estimators in (d) and (g) and (h). Who is smallest?

2. If gene frequencies are in equilibrium, the genotypes AA, Aa, and aa occur with probabilities $p_1 = (1-\theta)^2$, $p_2 = 2\theta(1-\theta)$, and $p_3 = \theta^2$, respectively. Plato et. al. (1964) published the following data on haptoglobin type in a sample of 190 people:

Haptoglobin Type		
$Hp1-1$	$Hp1-2$	$Hp2-2$
10	68	112

(a). Find the MLE of θ.
(b). Find asymptotic variance of the MLE.
(c). Find an approximate 95% confidence interval for θ.
(d). Carry out a Wald test of the hypothesis $H_0 : \theta = 3/4$ versus $H_1 : \theta \neq 3/4$ at the 5 percent level.
(e). Find the uinformly most powerful test of the hypothesis $H_0 : \theta = 3/4$ versus $H_1 : \theta > 3/4$ at the 5 percent level.
(f). Carry out a likelihood ratio test of the hypothesis $H_0 : \theta = 1/2$ versus $H_1 : \theta \neq 1/2$ at the 5 percent level. (This question is from The Rice book.)
(g). Carry out a goodness-of-fit test to check whether the data come from the assumed model.

3. The height of a child is not stable but increases over time. Since the pattern of growth varies from child to child, one way to understand the general growth pattern is by using the average of several children's heights, as presented in this data set. (Mean heights of a group of children in Kalama, an Egyptian village that is the site of a study of nutrition in developing countries. The data were obtained by measuring the heights of all 161 children in the village each month over several years.)

Variable Names:
1. Age: Age in months
2. Height: Mean height in centimeters for children at this age
The Data:

age	height
17	65.1
18	76.1
19	77
20	78.1
21	78.2
22	78.8
23	69.7
24	79.9
25	81.1
26	81.2
27	81.8
28	82.8
29	83.5

(a). Plot height versus age. Does the relationship appear linear?
(b). Fit the height as a linear function of age. Estimate the parameters and their standard errors.
(c). Find approximate 95% confidence intervals for the parameters.
(c). Test $H_0 : slope = 0$ versus $H_1 : slope \neq 0$ at the 5 percent level.
(d). Plot residuals versus age.
(e). Is the first observation (17, 66.5) a statistically plausible outcome?